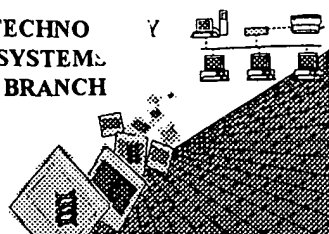


BIOTECHNO
SYSTEMS
BRANCH



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/937,187
Source: PT/09
Date Processed by STIC: 2/6/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER** **VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/efb/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name,
Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two,
2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 09/937/187

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.



PCT/09

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/937,187

DATE: 02/06/2002
TIME: 18:41:31

Input Set : A:\Neb-164.app
Output Set: N:\CRF3\02062002\I937187.raw

pp 1-5
**Does Not Comply
Corrected Diskette Needed**

OK 3 <110> APPLICANT: SANDMAN, KAREN E.
4 NOREN, CHRISTOPHER J.
5 NEW ENGLAND BIOLABS, INC.
7 <120> TITLE OF INVENTION: SURFACE DISPLAY OF SELENOCYSTEINE-CONTAINING PEPTIDES
9 <130> FILE REFERENCE: NEB-164-PCT
OK 11 <140> CURRENT APPLICATION NUMBER: US/09/937,187
12 <141> CURRENT FILING DATE: 2002-01-08
14 <150> PRIOR APPLICATION NUMBER: 60/134,286
15 <151> PRIOR FILING DATE: 1999-05-14
17 <160> NUMBER OF SEQ ID NOS: 42
19 <170> SOFTWARE: PatentIn Ver. 2.0
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 38
23 <212> TYPE: RNA
24 <213> ORGANISM: Synthetic *see item 10 on Erra summary sheet*
26 <220> FEATURE:
27 <223> OTHER INFORMATION: N = A, G, C, or U
29 <220> FEATURE:
30 <223> OTHER INFORMATION: K = G and U
OK 32 <400> SEQUENCE: 1
33 nnknnknnkn nkugannknn knnkucggcc gaaacaug 38
35 <210> SEQ ID NO: 2
36 <211> LENGTH: 24
37 <212> TYPE: DNA
38 <213> ORGANISM: Synthetic *item 10*
40 <400> SEQUENCE: 2
41 tcgtcttttc cttgaaagtc gcct 24
43 <210> SEQ ID NO: 3
44 <211> LENGTH: 24
45 <212> TYPE: DNA
46 <213> ORGANISM: Synthetic
48 <400> SEQUENCE: 3
49 aagtgtacgc tttgatctat gctg 24
51 <210> SEQ ID NO: 4
52 <211> LENGTH: 24
53 <212> TYPE: DNA
54 <213> ORGANISM: Synthetic
56 <400> SEQUENCE: 4
57 ttgcttttgc cttgaaatgt tctt 24
59 <210> SEQ ID NO: 5
60 <211> LENGTH: 24
61 <212> TYPE: DNA
62 <213> ORGANISM: Synthetic

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/937,187

DATE: 02/06/2002

TIME: 18:41:31

Input Set : A:\Neb-164.app

Output Set: N:\CRF3\02062002\I937187.raw

64 <400> SEQUENCE: 5
65 atgactacgc agtgaccttc tctg 24
67 <210> SEQ ID NO: 6
68 <211> LENGTH: 24
69 <212> TYPE: DNA
70 <213> ORGANISM: Synthetic
72 <400> SEQUENCE: 6
73 catattccgc cgtgaacgaa tcct 24
75 <210> SEQ ID NO: 7
76 <211> LENGTH: 24
77 <212> TYPE: DNA
78 <213> ORGANISM: Synthetic
80 <400> SEQUENCE: 7
81 aaggctctgt gttgacagga ttcg 24
83 <210> SEQ ID NO: 8
84 <211> LENGTH: 24
85 <212> TYPE: DNA
86 <213> ORGANISM: Synthetic
88 <400> SEQUENCE: 8
89 cttcttcogt gttgagctca gccg 24
91 <210> SEQ ID NO: 9
92 <211> LENGTH: 24
93 <212> TYPE: DNA
94 <213> ORGANISM: Synthetic
96 <400> SEQUENCE: 9
97 catcatccga cttgagctaa gcag 24
99 <210> SEQ ID NO: 10
100 <211> LENGTH: 24
101 <212> TYPE: DNA
102 <213> ORGANISM: Synthetic
104 <400> SEQUENCE: 10
105 atgcctccta cgtgaatggc tacg 24
107 <210> SEQ ID NO: 11
108 <211> LENGTH: 24
109 <212> TYPE: DNA
110 <213> ORGANISM: Synthetic
112 <400> SEQUENCE: 11
113 aattggtttt cttgactgac tacg 24
115 <210> SEQ ID NO: 12
116 <211> LENGTH: 24
117 <212> TYPE: DNA
118 <213> ORGANISM: Synthetic
120 <400> SEQUENCE: 12
121 ctgcatccga cgtgagctcg gcct 24
123 <210> SEQ ID NO: 13
124 <211> LENGTH: 24
125 <212> TYPE: DNA
126 <213> ORGANISM: Synthetic
128 <400> SEQUENCE: 13

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/937,187

DATE: 02/06/2002
TIME: 18:41:31

Input Set : A:\Neb-164.app
Output Set: N:\CRF3\02062002\I937187.raw

129 gatagggggc cttgagcgaa gatt	24
131 <210> SEQ ID NO: 14	
132 <211> LENGTH: 24	
133 <212> TYPE: DNA	
134 <213> ORGANISM: Synthetic	
136 <400> SEQUENCE: 14	
137 gcgtctttgc cttgaaggac gagt	24
139 <210> SEQ ID NO: 15	
140 <211> LENGTH: 24	
141 <212> TYPE: DNA	
142 <213> ORGANISM: Synthetic	
144 <400> SEQUENCE: 15	
145 ttgccgcgtc agtgatggtc tccg	24
147 <210> SEQ ID NO: 16	
148 <211> LENGTH: 24	
149 <212> TYPE: DNA	
150 <213> ORGANISM: Synthetic	
152 <400> SEQUENCE: 16	
153 ttgactggta cgtgatgtca gaat	24
155 <210> SEQ ID NO: 17	
156 <211> LENGTH: 24	
157 <212> TYPE: DNA	
158 <213> ORGANISM: Synthetic	
160 <400> SEQUENCE: 17	
161 gaggcgtcgc gttgatgttc gact	24
163 <210> SEQ ID NO: 18	
164 <211> LENGTH: 24	
165 <212> TYPE: DNA	
166 <213> ORGANISM: Synthetic	
168 <400> SEQUENCE: 18	
169 aagttggctc gttgatcggc gtcg	24
171 <210> SEQ ID NO: 19	
172 <211> LENGTH: 24	
173 <212> TYPE: DNA	
174 <213> ORGANISM: Synthetic	
176 <400> SEQUENCE: 19	
177 aatggggcgc agtgatcgag gcat	24
179 <210> SEQ ID NO: 20	
180 <211> LENGTH: 24	
181 <212> TYPE: DNA	
182 <213> ORGANISM: Synthetic	
184 <400> SEQUENCE: 20	
185 gcgagtccta cttgatttaa gccg	24
187 <210> SEQ ID NO: 21	
188 <211> LENGTH: 24	
189 <212> TYPE: DNA	
190 <213> ORGANISM: Synthetic	
192 <400> SEQUENCE: 21	
193 tgtgtcctc cgtgatctac tcgt	24

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/937,187

DATE: 02/06/2002
 TIME: 18:41:31

Input Set : A:\Neb-164.app
 Output Set: N:\CRF3\02062002\I937187.raw

```

195 <210> SEQ ID NO: 22
196 <211> LENGTH: 24
197 <212> TYPE: DNA
198 <213> ORGANISM: Synthetic
200 <400> SEQUENCE: 22
201 cagtcgacgc ggtgatggaa tgat 24
203 <210> SEQ ID NO: 23
204 <211> LENGTH: 24
205 <212> TYPE: DNA
206 <213> ORGANISM: Synthetic
208 <400> SEQUENCE: 23
209 attgtggagt cgtgattgaa tccg 24
211 <210> SEQ ID NO: 24
212 <211> LENGTH: 24
213 <212> TYPE: DNA
214 <213> ORGANISM: Synthetic
216 <400> SEQUENCE: 24
217 acgcagcgta tgtgattgcc gccc 24
219 <210> SEQ ID NO: 25
220 <211> LENGTH: 24
221 <212> TYPE: DNA
222 <213> ORGANISM: Synthetic
224 <400> SEQUENCE: 25
225 gtgcagtata cgtgattgcc gaag 24
227 <210> SEQ ID NO: 26
228 <211> LENGTH: 24
229 <212> TYPE: DNA
230 <213> ORGANISM: Synthetic
232 <400> SEQUENCE: 26
233 gctgggcagt cgtgatcgac tgat 24
235 <210> SEQ ID NO: 27
236 <211> LENGTH: 24
237 <212> TYPE: DNA
238 <213> ORGANISM: Synthetic
240 <400> SEQUENCE: 27
241 ctgtctgcga gtcgatcgca gttt 24
243 <210> SEQ ID NO: 28
244 <211> LENGTH: 8
245 <212> TYPE: PRT
246 <213> ORGANISM: Synthetic
248 <220> FEATURE:
249 <223> OTHER INFORMATION: At position 5, X = Selenocysteine
251 <400> SEQUENCE: 28
252 Ser Ala Arg Val Xaa His Gly Pro
253 1 5
256 <210> SEQ ID NO: 29
257 <211> LENGTH: 98
258 <212> TYPE: DNA
259 <213> ORGANISM: Synthetic
  
```

OK->

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/937,187

DATE: 02/06/2002

TIME: 18:41:31

Input Set : A:\Neb-164.app

Output Set: N:\CRF3\02062002\I937187.raw

261 <400> SEQUENCE: 29
 262 catgtttcgg cgtaccgac cgattggtgc agacctgcaa ccgatggggcc gtgtcagaca 60
 263 cgagcgctag agtgagaata gaaaggtacc cgggcatg 98
 265 <210> SEQ ID NO: 30
 266 <211> LENGTH: 25
 267 <212> TYPE: DNA
 268 <213> ORGANISM: Synthetic
 270 <400> SEQUENCE: 30
 271 catgcccggg tacctttcta ttctc 25
 273 <210> SEQ ID NO: 31
 274 <211> LENGTH: 20
 275 <212> TYPE: DNA
 276 <213> ORGANISM: Synthetic
 278 <400> SEQUENCE: 31
 279 ccctcatagt tagcgtaacg 20
 281 <210> SEQ ID NO: 32
 282 <211> LENGTH: 10
 283 <212> TYPE: PRT
 284 <213> ORGANISM: Synthetic
 286 <220> FEATURE:
 287 <223> OTHER INFORMATION: At position 5, X = Selenocysteine
 289 <400> SEQUENCE: 32
 OK 290 Ser Ala Arg Val Xaa His Gly Pro Ser Val
 291 1 5 10
 294 <210> SEQ ID NO: 33
 295 <211> LENGTH: 85
 296 <212> TYPE: DNA
 297 <213> ORGANISM: Synthetic
 299 <220> FEATURE:
 300 <223> OTHER INFORMATION: M = A or C
 302 <220> FEATURE:
 303 <223> OTHER INFORMATION: N = A, C, T or G
 305 <400> SEQUENCE: 33
 OK 306 catgtttcgg cgtattggtg cagacctgca accgamnnmn nmntcamnn mnnmnnmna 60
 307 gagtgagaat agaaaggtac ccggg 85
 309 <210> SEQ ID NO: 34
 310 <211> LENGTH: 85
 311 <212> TYPE: DNA
 312 <213> ORGANISM: Synthetic
 314 <220> FEATURE:
 315 <223> OTHER INFORMATION: M = A or C
 317 <220> FEATURE:
 318 <223> OTHER INFORMATION: N = A, C, T or G
 320 <400> SEQUENCE: 34
 OK 321 catgtttcgg cgtattggtg cagacctgca accgamnnmn nmnatcamnn mnnmnnmna 60
 322 gagtgagaat agaaaggtac ccggg 85
 324 <210> SEQ ID NO: 35
 325 <211> LENGTH: 8
 326 <212> TYPE: PRT

Use of n and/or Xaa has been detected in the Sequence Listing.
 Review the Sequence Listing to insure a corresponding
 explanation is presented in the <220> to <223> fields of
 each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/937,187

DATE: 02/06/2002

TIME: 18:41:32

Input Set : A:\Neb-164.app

Output Set: N:\CRF3\02062002\I937187.raw

L:11 M:270 C: Current Application Number differs, Replaced Application Number
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:33 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:1
L:33 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:1
L:33 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:252 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:28
L:252 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:28
L:252 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28
L:290 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:32
L:290 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:32
L:290 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32
L:306 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:33
L:306 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:33
L:306 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33
L:321 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:34
L:321 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:34
L:321 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:34
L:383 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:40
L:383 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:40
L:383 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40
L:397 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:41
L:397 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:41
L:397 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41